Research Article

Received: 23 January 2024 Revised: 15 March 2024 Accepted article published: 31 March 2024 Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI 10.1002/ps.8115

First use of unmanned aerial vehicles to monitor *Halyomorpha halys* and recognize it using artificial intelligence

Daniele Giannetti,[a,b†*] Niccolò Patelli,[a†] Lorenzo Palazzetti,[c†] Francesco Betti Sorbelli,[c] Cristina M. Pinotti[c‡] and Lara Maistrello[a,d‡*]

Abstract

BACKGROUND: *Halyomorpha halys* is one of the most damaging invasive agricultural pests in North America and southern Europe. It is commonly monitored using pheromone traps, which are not very effective because few bugs are caught and some escape and/or remain outside the trap on surrounding plants where they feed, increasing the damage. Other monitoring techniques are based on visual sampling, sweep-netting and tree-beating. However, all these methods require several hours of human labor and are difficult to apply to large areas. The aim of this work is to develop an automated monitoring system that integrates image acquisition through the use of drones with *H. halys* detection through the use of artificial intelligence (AI).

RESULTS: The study results allowed the development of an automated flight protocol using a mobile app to capture high-resolution images. The drone caused only low levels of disturbance in both adult and intermediate instars, inducing freezing behavior in adults. Each of the AI models used achieved very good performance, with a detection accuracy of up to 97% and recall of up to 87% for the X-TL model.

CONCLUSION: The first application of this novel monitoring system demonstrated the potential of drones and AI to detect and quantify the presence of *H. halys*. The ability to capture high-altitude, high-resolution images makes this method potentially suitable for use with a range of crops and pests.

Supporting information may be found in the online version of this article.

Keywords: brown marmorated stink bug; drones; freezing behavior; automated pest monitoring; integrated pest management

1 INTRODUCTION

Unmanned aerial vehicles (UAVs or drones) are robots that are remotely or autonomously flown by human operators.[1] Because of their handling, speed and hovering capabilities, UAVs are receiving increasing attention and being used more widely in environmental sciences,[2,3] wildlife monitoring[4–6] and agroecosystems.[7–9] However, the use of drones for arthropod monitoring has received little attention. UAVs have been used to identify immobile instars like *Monema flavescens* chrysalis, a major pest of urban trees in east Asia.[10] Thanks to new high-quality cameras, it is possible to obtain images that can be useful for monitoring mobile invertebrate instars, such as adults of the brown marmorated stink bug *Halyomorpha halys* (Hemiptera, Pentatomidae). Native to Asia, *H. halys* has become one of the most damaging invasive agricultural pests in North America and Europe. First officially detected in 2012,[11] it quickly spread and became a key pest of orchards in northern Italy,[12] causing losses of up to 80–100% and an estimated € 588 million in damage to fruit production in 2019 (CSO, 2021[13]). *Halyomorpha halys* is commonly monitored using traps baited with aggregation pheromones that can attract stink bugs from up to 70 m when used in association with host plants within an orchard.[14] The arrestment area is limited to a radius of 2.5 m around the bait, regardless of the concentration of aggregation pheromone used.[15] However, these traps are not very effective at capturing individuals; few are caught, and some escape and/or remain outside the trap on surrounding plants where they can cause damage (trap spillover).[16] Other monitoring techniques include counting individuals during standardized visual sampling, sweep-netting (on herbaceous crops) and tree-

* Correspondence to: D Giannetti or L Maistrello, Department of Life Sciences, University of Modena and Reggio Emilia, Reggio Emilia, Italy, E-mail: daniele.giannetti@unipr.it (Giannetti); lara.maistrello@unimore.it (Maistrello)

† Co-first authors. ‡ Co-last authors.

a Department of Life Sciences, University of Modena and Reggio Emilia, Reggio Emilia, Italy

b Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy

c Department of Computer Science and Mathematics, University of Perugia, Perugia, Italy

d NBFC, National Biodiversity Future Center, Palermo, Italy

beating techniques.[12] However, all these methods require several hours of human labor and are difficult to apply to large areas. To minimize yield losses and maximize efficiency, a reliable monitoring technique that does not involve pheromones or hours of labor needs is required. One such example is the use of UAVs. Significant progress has been made in optimizing trajectory planning for autonomous drones. Recent research[17,18] has demonstrated the effectiveness and efficiency of their theoretical algorithms. Unlike the current research, these monitoring strategies require advanced knowledge of the pest's spatial distribution, making them quite useless in practice. Moreover, one area of current research is the development of machine learning (ML)-oriented solutions for monitoring insect species. Traditional ML methods, such as support vector machines and adaptive boosting have shown promise in automation of the scouting process.[19,20] However, the introduction of convolutional neural networks (CNNs) represents a breakthrough in model detection of invasive species.[21,22] Indeed, the extreme capacity of CNNs to recognize and infer complex patterns in images represents a pivotal moment in enhancing integrated pest management strategies and reducing chemical usage by pre-emptive pest detection. For example, an innovative approach to the early detection and continuous monitoring of whiteflies (*Bemisia tabaci*) and thrips (F*rankliniella occidentalis*) in greenhouses employs an image-processing algorithm and artificial neural networks.[23] CNNs have now been used to monitor *H. halys* with promising results.[24,25] This has been investigated by employing several CNNs and learning procedures on data acquired in the laboratory and mostly relying on a human operator. Despite the notable performance achieved in this research, the applicability of CNNs has been confined to a controlled environment, and so is not directly applicable in an orchard. The aim of our work is to test the use of commercial drones for the automated identification and monitoring of *H. halys* in pear orchard agroecosystems, as part of the 'HALY.ID' project.[26] Specifically, we aimed to: (i) evaluate the influence of the drone on *H. halys* behavior; (ii) compare monitoring performed using the drone with monitoring by an operator; (iii) evaluate the quality of the images obtained for artificial intelligence (AI)-based identification; and (iv) test automatic recognition of *H. halys* using AI.

2 MATERIALS AND METHODS

2.1 Study area

The experiment was conducted in Emilia Romagna near the city of Carpi, Modena, Italy (GPS 44.729915 N, 10.874830 E) in an area planted with Abate Fetel pears (field size: 40 m × 32 m, 176 plants). Five pheromone baits (PHEROCON BMSB DUAL™, Trécé Inc.) were randomly placed in the field to increase the presence of *H. halys*. Each bait contained both the two-component aggregation pheromone (3*S*,6*S*,7*R*,10*S*)-10,11-epoxy-1-bisabolen-3-ol and (3*R*,6*S*,7*R*,10*S*)- 10,11-epoxy-1-bisabolen-3-ol,[27] and the pheromone synergist (methyl (2*E*,4*E*,6*Z*)-decatrienoate).[28]

2.2 UAV selection

The HALY.ID[29] project required careful hardware selection to meet the constraints imposed by the data-collection task. One constraint was that regulations required the use of off-the-shelf hardware. In addition, the objective of the study was to use computer vision techniques to detect *H. halys* in an orchard, and this requires a high-resolution camera because the smallest feature of *H. halys* is 0.2 mm in size. However, drones equipped with cameras of a suitable resolution faced operational limitations owing to their obstacle avoidance systems. Therefore, the best option was to select a DJI Zenmuse H20 camera for its ability to capture images at different distances while maintaining the desired resolution, in combination with the DJI Matrice 300 (6.3 kg), which has a more reliable collision avoidance system. Moreover, the Matrice 300 is equipped with real-time kinematic positioning, which allows for extremely accurate global positioning from multiple satellites,[30] with a margin of error as low as 1 cm when connected to between 12 and 20 satellites (numbers obtained from the drone's radio command). The main downside of this equipment is the bulkiness of the drone, which with its propellers open has a diameter of ∼1 m. Therefore, the minimum width of a single aisle in an orchard needs to be ≥3 m, considering both the size of the UAV and the collision avoidance system. Currently, aisles are almost never ≥3 m; however, in view of the promising results obtained in the flight protocol used here, which requires the UAV to fly at 10 m, a height ∼6 m above the top of the canopy does not represent a limitation. The DJI Zenmuse H20 camera (∼ 0.7 kg) is a stabilized system with a 12 Mega Pixel (MP) wide angle, and the 20 MP sensor has a 23× optical zoom. To capture details of the body of *H. halys*, we aimed for a resolution distance (σ) of 0.2 mm (*H. halys* smallest feature). The maximum diagonal field of view (F_D) for discerning these features is determined by the product of the number of pixels on the chip's diagonal and the resolution distance; that is, $0.2 \times 6483 = 1296$ mm. Using the chip diagonal angle of view, the diagonal dimension of the solid angle at the lens opposite the Diagonal Field of View (DFoV), we can correlate the DFoV with the distance from which the subject is captured the diagonal angle of view (α_d). Fixing α_d and F_D, one finds that $\mu = F_D / 2\tan(\alpha_d/2)$ can be used to capture an image complying with the desired F_D. Because the DAoV of the DJI Zenmuse H20 varies between 4° and 66.6°, we can achieve the desired $F_D = 1296$ mm by selecting any distance between 0.98 m and 18.5 m. As a result, we found that the optimal drone altitude is 10.5 m, which simultaneously ensures a minimum distance of ∼10.6 m between the drone and the tree canopy, and a risk-free flight without any obstacles.

2.3 Insect rearing

Adult *H. halys* were collected in the field from urban parks in the provinces of Modena, Reggio Emilia and Parma (Emilia Romagna, Italy) using a tree-beating technique. Stink bugs were reared in BugDorm cages (32.5 cm × 32.5 cm × 32.5 cm) and placed in climatic chambers at 26 °C, 60% relative humidity and a 16:8 h light/dark photoperiod. Each cage contained up to 50 adults in a sex ratio of 50:50. Stink bugs were fed with fresh organic fruits (carrots, apples, tomatoes, green beans) and peanuts twice a week. In addition, wet cotton wicks were added to each cage as a water source.[31,32]

2.4 Effects of UAV on *H. halys* behavior

A field experiment to quantify the effects of drone use on *H. halys* behavior was conducted for all pre-imaginal instars ($n = 40$ for each instar from the first to fifth nymphal instar) and adult insects (males, $n = 40$; females, $n = 40$). Each insect (one nymph for stage) was placed on a pear tree at the end of a branch that was at least 30 cm long, at a height of 1 m above the ground and was allowed to settle for 1 min before the drone was positioned on hovering flight. After positioning the drone, continuous sampling was conducted for 1 min, recording the insect's behavioral response in terms of: (i) walking; (ii) falling

(from the leaf or branch); (iii) hiding behind leaves or branches; (iv) freezing (the insect suddenly stops all activity, remains motionless and grasps the substrate firmly with its legs)[33]; and (v) flying (adults only). Three flight heights (1, 4 and 8 m) were selected (Fig. 1), with the drone hovering perpendicular to the branch for 2 min at each height. The intensity of the wind generated by the hovering drone wad previously measured by placing a wireless anemometer at the end of the branch where the insect was placed. Specifically, wind intensity was 1 m/s at 1 m, 4 m/s at 4 m and 2 m/s at 8 m. Control experiments were conducted by inducing manual disturbance (an operator moving the branch with forceps) where the *H. halys* individuals were positioned, after the 1-min settling period. Field experiments were conducted from 7 January to 20 July 2022 (10:00 to 12:00 am). A chi-squared test and analysis of standardized residuals were used to evaluate the effects of UAV on behavior.

2.5 Effects of wind and noise on *H. halys* behavior

Specific tests were conducted to separately evaluate the effects of wind and drone noise on adult behavior. To create wind without noise, a set of silent plastic computer cooling fans ranging in diameter from 6 to 20 cm was set up, with a flow controller adjusted to allow for different fan intensities measured at the point where the insect was positioned. A total of 20 adults (10 males and 10 females) were used for wind intensities of: (i) 0.5 m/s, (ii) 1 m/s, and (iii) 1.5 m/s (Fig. 2(a)). To generate noise without wind, Bluetooth speakers (Fig. 2 (b)) were used to reproduce the sound generated by the drone propeller during flight (86 dB). The sound had previously been measured and recorded using a sound level meter. In total 20 adults (10 males and 10 females) were used. The results were analyzed with a chi-squared test and analysis of standardized residuals.

2.6 Flight protocol for image acquisition and image quality analysis

Nine sampling points were selected in each field (three per row in three different rows), two on opposite sides of the field and one in the middle. The drone spent an average of 2 min 30 s at each of sampling point (Fig. 3). Unlike previous work,[34,35] in which the authors experienced unexpected blockages because of the limited space inside the aisles, a flight protocol with an altitude of 10 m was used to avoid this problem and any possible collisions. At each sampling position, the UAV system was set to build a mosaic-like composition of 20 photos taken using the ×15 zoom of the DJI Zenmuse H20. The photo mosaic is divided into five columns and four rows, which allows coverage of an area of 2.8 m height and 5.75 m width; each photo covers an area of 0.7 m × 1.15 m. To ensure that the accuracy was the same as with operator's visual sampling, the area was marked with a red and white band at the boundary. Flights were conducted from 24 August to 28 September 2022 for a total of 16 missions (from 08:30 am to 12:30 pm). Only images containing the adult stage of *H. halys* were selected. To assess image quality in the context of AI applications, we chose the metrics proposed by Narvekar and Ware[36,37] to assess blurriness and brightness. The first metric uses a no-reference image blur estimator, which is inspired by human blur perception at different contrast levels. It uses a probabilistic model to estimate the likelihood of detecting blurriness at individual edges in an image. The probabilities are then aggregated by computing the cumulative probability of blur detection. The second metric uses a nonlinear function to approximate the human visual response curve. This is achieved by transforming the color channels into the luminance color space.

2.7 Identification using AI (algorithm selection, training, validation and automatic detection test)

YOLO[38,39] is a deep learning-based architecture built on top of PyTorch. It redefines object detection as a regression problem,

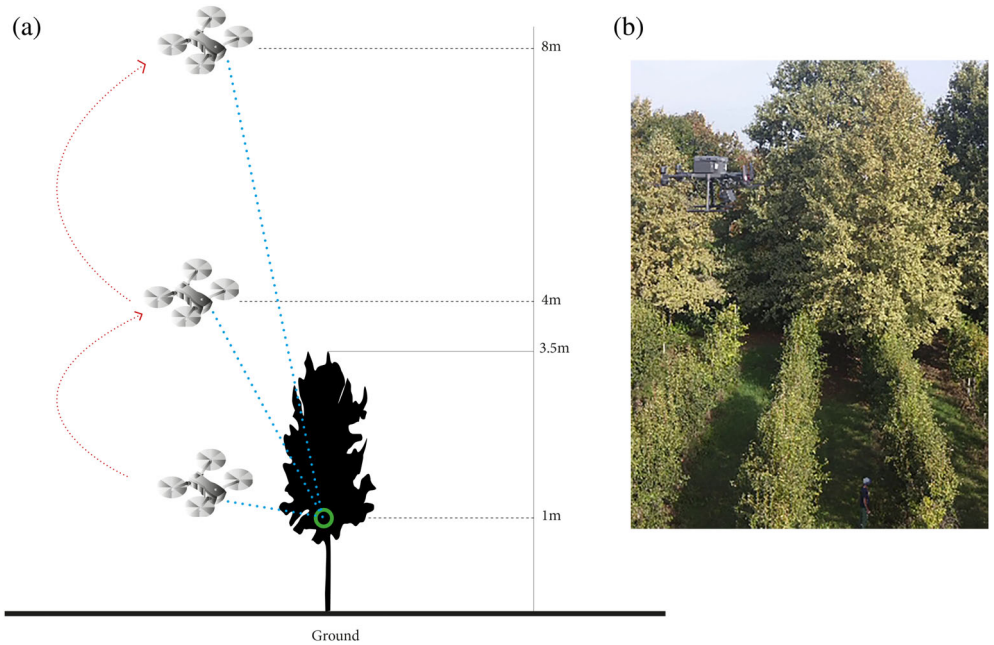

Figure 1. (a) Scheme of different flight altitudes tested (1, 4 and 8 m). The green circle at 1 m represents the spot where the specimens were positioned. (b) View of the drone at 8 m (upper left-hand corner) compared with an operator in the field (lower right-hand corner).

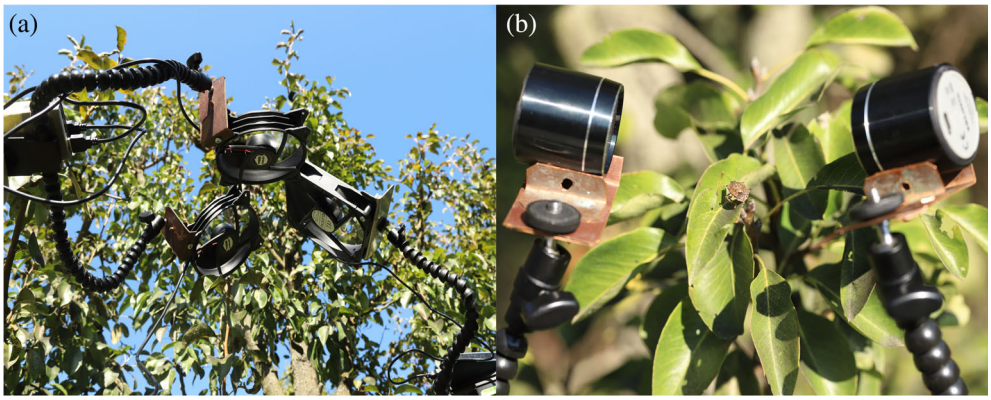

Figure 2. (a) Silent computer fan used to reproduce wind without noise (wind strength was measured with an anemometer beforehand). (b) Bluetooth computer speakers used to reproduce the noise without wind (86 dB).

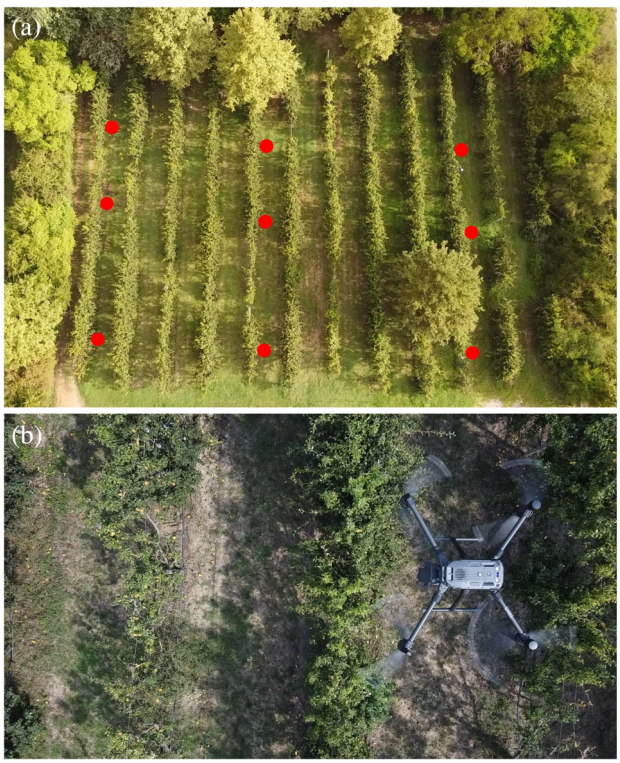

Figure 3. (a) The nine sampling points in the experimental field. (b) View of the Matrice 300 hovering over the experimental field during a sampling mission.

spatially separating bounding boxes and assigning probabilities using a single CNN. Notably, YOLO is highly efficient, providing a lightweight and fast solution that requires fewer computational resources than other state-of-the-art models, while maintaining competitive performance.[39] It can process images at a remarkable rate of 45 frames per second, making it ideal for real-time bug detection, even on edge computing devices within a drone. In addition, YOLO offers five different models with different scales, each characterized by an increasing number of hyper-parameters, such as the number of layers: nano (N), small (S), medium (M), large (L) and extra-large (X). Larger models typically have more layers and deeper architectures, allowing them to capture more complex patterns and features. Moreover, larger models perform better on complex data sets with small objects or intricate details. However, smaller models may be more suitable for real-time applications or resource-constrained environments. In the following, the models M and X are used for our bug detection task; we used the standard learning parameters suggested by the developers of YOLO.[38,39] Because of limited image data and computational resources, we chose to train our model either from scratch or using transfer learning (TL).[40] TL is a valuable computer vision technique that uses models pretrained on large image data sets. This knowledge is applied to the new model, even when only a small number of images are available for the new task. TL involves selecting and transferring relevant components from an existing model to address a related problem. Central to TL is the concept of generalization—only knowledge applicable to different scenarios or conditions is transferred. Unlike models that are built from scratch, TL-based models are versatile and adaptable across different scenarios and data sets. This adaptability, combined with the ability to extract valuable insights from a pretrained model, is crucial in efficiently and effectively identifying stink bugs in our research. Each image was reviewed to evaluate the presence of *H. halys*. In addition, labeling was performed by an expert human operator using the well-known open-source cloud service makesense.ai (P Skalski, Make Sense. https://github.com/SkalskiP/make-sense/, 2019). It should be noted that the labeling process consists of outlining instances of the bug with a rectangle. Despite the bug's limited dimensions, we decided to include feet and antennae in the annotations. The motivation behind this decision is that because of the acquisition protocol feet and antennae remain visible (Fig. 4). After careful annotation of the *H. halys* instances inside the images we revised the data set's suitability for computer vision. First, we evaluated the size of the bounding boxes (the size of the *H. halys* captured) and found that they exhibited limited size variation, ranging from 33×45 to 183×132 pixels. This uniformity stemmed from our controlled acquisition process and careful calibration of the camera settings. Second, we visualized the *H. halys* relative coordinates (the x and y components of bounding boxes' center), defined as a percentage with respect to the image size. In this case, we observed significant variability in their positions, reflecting the unpredictable nature of bug positions during drone flights. It should be noted that both the size and position distribution were evaluated on all images with at least one *H. halys* inside the data set. This diversity ensured that the data set encompassed a wide range of bug positions and poses (Fig. 5) paving the way

Figure 4. (a) Image captured by the drone. (b,c) Illustration of image resolution (b) and *Halyomorpha halys* annotations (c).

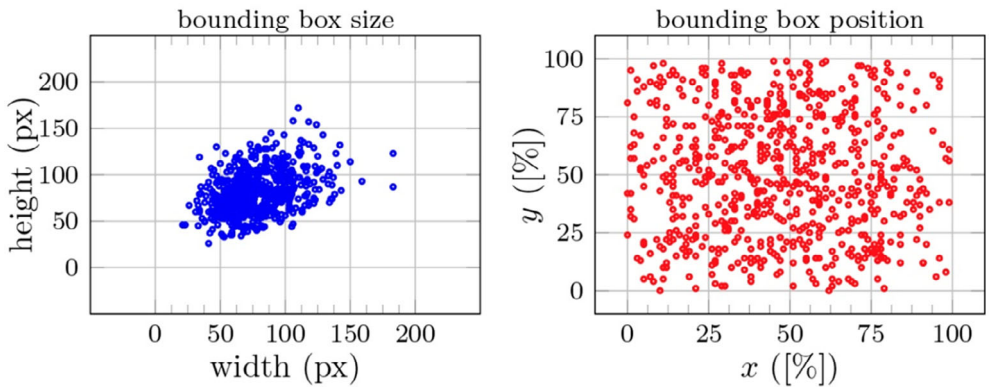

Figure 5. Distribution of bounding box size (left) and position (right). The majority of bounding boxes exhibit a size of ∼100 × 100 pixels (left), with their positions scattered randomly throughout the images (right).

for a promising data set suitable for an insect detection task. In our simulation, we used a rigorous evaluation framework to assess the performance of our object detection models. Central to this evaluation are fundamental metrics including intersection over union (IoU), true positive (TP), false positive (FP) and false negative (FN). These metrics serve as foundational components for deriving more comprehensive measures such as precision, recall, mean average precision (mAP) for the Pascal Visual Object Classes (VOC) data set and mAP for the COCO data set. Table 1 gives a formal definition of the metrics used.

3 RESULTS

3.1 Effects of UAV on *H. Halys* behavior

The results show significantly different responses for instar N1 in all behaviors, highlighting an operator effect with a decrease in the number of individuals walking ($\chi^2 = 71.31$, df = 3, $P < 0.001$) and an increase in those falling ($\chi^2 = 24.12$, df = 3, $P < 0.001$) and hiding ($\chi^2 = 30.31$, df = 3, $P < 0.001$) (Fig. 6(a)). The analyses did not show any disturbance effects by the drone as a function of flight height, but do show a high disturbance effect of the operator even for instars N2, N3, N4 and N5. For instar N2, a significant reduction in the number of individuals walking ($\chi^2 = 42.20$, df = 3, $P < 0.001$) and an increase in the number of those falling ($\chi^2 = 31.82$, df = 3, $P < 0.001$) were observed, although no differences were detected for hiding ($\chi^2 = 11.63$, df = 3, $P = 0.008$) (Fig. 6(b)). Similarly, for N3, there was a statistically significant effect with an operator-induced increase in the number of individuals falling ($\chi^2 = 92.17$, df = 3, $P < 0.001$) and fewer walking ($\chi^2 = 61.31$, df = 3, $P < 0.001$) (Fig. 6(c)). For instar N4, we also recorded an operator effect with a highly significant difference in the number of individuals falling ($\chi^2 = 88.17$,

Table 1. Summary of the employed metrics

Metric	Definition	Formula
Intersection over Union (IoU)	Measures overlap between predicted and ground truth bounding boxes	IoU = (Intersection Area) / (Union Area)
True Positive (TP)	Correctly predicted objects	—
False Positive (FP)	Incorrectly predicted objects	—
False Negative (FN)	Missed detections	—
Precision (P)	Accuracy of positive predictions	TP / (TP + FP)
Recall (R)	Model's ability to identify relevant instances	TP / (TP + FN)
mAP Pascal ($mAP_{0.5}$)	Average precision across classes with different confidence levels	(1 / n_classes) × Σ(AP_i)
mAP COCO ($mAP_{0.95}$)	Averaged mean average precision Pascal across different IoU thresholds	(1 / n_IOU_thresholds) × Σ(mAP_i)

df = 3, $P < 0.001$) and, as above, fewer walking ($\chi^2 = 62.67$, df = 3, $P < 0.001$), but no differences in the number of individuals hiding ($\chi^2 = 0.40$, df = 3, $P = 0.93$) (Fig. 1(d)). For instar N5, the analysis shows significant differences between treatments as in previous stages, with an operator effect on the number of individuals falling ($\chi^2 = 45.00$, df = 3, $P < 0.001$) and walking ($\chi^2 = 18.64$, df = 3, $P = 0.003$), but no differences in the number hiding ($\chi^2 = 1.87$, df = 3, $P = 0.59$). Moreover, field observations show that N5 exhibited freezing behavior in the presence of the drone at 4 m ($n = 3$) and 8 m ($n = 5$) (Fig. 1(e)).

For adults, the presence of the drone at 4 and 8 m did not affect falling or hiding behavior, although a significant increase in walking was only observed at 1 m (Fig. 7). However, the presence of the drone at 4 and 8 m significantly increased the number of individuals exhibiting freezing behavior (85% of insects tested), whereas the presence of the operator significantly increased flying behavior, exhibited by 87.5% of individuals tested (Fig. 7).

3.2 Effects of wind and noise on *H. halys* behavior

Considering the effects of wind and noise, a significant difference was detected between treatments ($\chi^2 = 42.20$, df = 3, $P < 0.001$). Specifically, no freezing behavior was observed at a wind speed of 0.5 m/s; on the contrary, freezing was exhibited by 45% of individuals tested at the speed of 1 m/s and by 85% individuals at the speed of 1.5 m/s. Finally, no freezing behavior was recorded in the sound tests.

3.3 Flight protocol for acquired images and image quality analysis

A total of 2459 images were acquired during the 16 flight missions in 2022. After careful analysis by an expert operator, it was determined that only 402 of these images featured *H. halys*, representing a total of 546 individuals, the majority of which were adults ($n = 528$), followed by the pre-imaginal instars N5 ($n = 11$), N4 ($n = 6$) and N3 ($n = 1$). The image quality assessment represents the respective metrics computed on our data set, performed only on images showing the presence of the adult stage of *H. halys*. It is worth noting that the drone captures images leveraging only RGB bands. Despite there being several promising bands available to detect the presence of *H. halys*, we believe that RGB cameras are more practical for autonomous in-field detection protocols because of their broader diffusion and minor cost. The plot on the left shows the blurriness levels of the collected images, whereas the plot on the right shows the achieved brightness (Fig. 8). In both plots, the images are displayed on the *x*-axis and the metric scores are plotted on the *y*-axis. The blur score falls in the range [0, 1], where 0 represents a completely blurred image, and 1 represents an extremely sharp one. Similarly, the brightness metric is a score ranging from 0 to 1, where 0 indicates no light and 1 indicates glaring light. Consequently, images with favorable metrics should have a blur score close to 1 and a brightness score around 0.5. We found that the blurriness scores averaged around 0.6, indicating sharp images, and the brightness scores remained stable at around 0.4, close to the optimal value of 0.5. Hence, the proposed acquisition protocol guarantees notable image quality stability, overcoming issues[34] related to data set heterogeneity. In addition, because the protocol acquires images relying on the drone camera using a constrained strategy, all the criticalities[41] were tackled. Indeed, the obtained data set homogeneity prevents use of an artificial data set, and simultaneously paves the way for a solid AI model for *H. halys* detection. The reasoning behind the latter observation is that the closer the data are to reality, the better the detection by the AI model.

3.4 Identification of *H. halys* using AI (algorithm selection, training, validation and automatic detection test)

The image data set with the presence of *H. halys* was partitioned as above: 70% was allocated for training, 20% for validation and 10% for testing. Thus, 282 images were used for training, 80 for validation and 40 for testing. We had a limited number of training samples, so we created more by generating three additional images for each training image through random transformations that included variations in hue, saturation and value, translation, scaling, flipping and mosaic techniques. The applied transformations comprise both geometrical and color space transformations that can be experienced during a flight. This process resulted in 1128 training images. It is important to note that training augmentation is commonly used to prevent models from overfitting during the training process and to make them more robust against small variations within images. However, our acquisition protocol is not susceptible to such variations, so further investigations are not required. In Fig. 9, we report the training performance of the best resulting models. Each plot shows the score on the *y*-axis and the epoch on the *x*-axis. The score represents either the value returned by the box loss (Fig. 9(a)) or the detection metrics results for the models (Fig. 9(b–d)). In Fig. 9(a) the box loss shows the model's ability to predict bounding box coordinates. By contrast, the number of epochs represents the number of iterations on which neural network's weights are adjusted. All the models rapidly reach a low box loss, although TL models converge faster. The remaining three plots display the metrics achieved by the models during the validation phase.

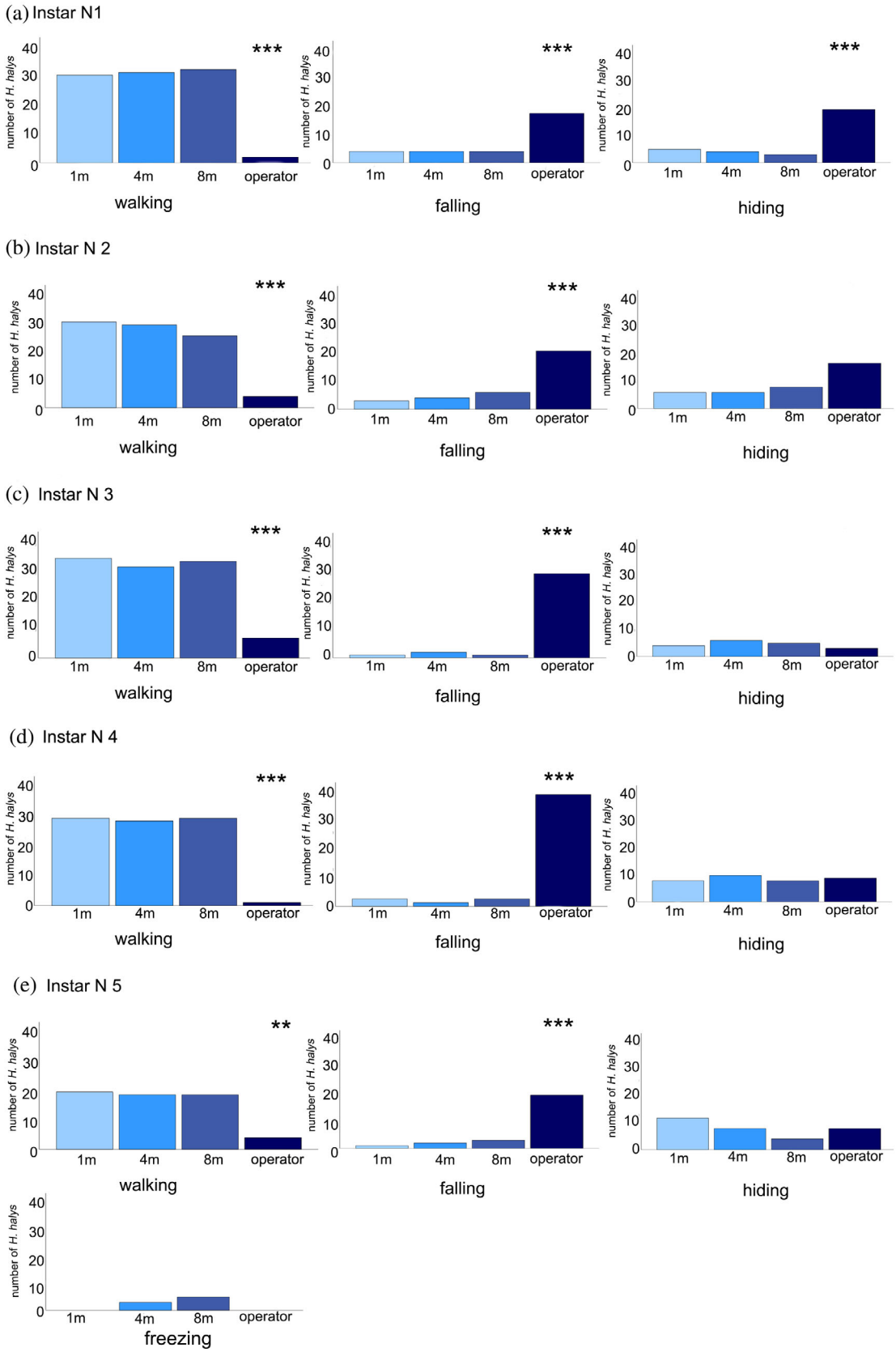

Figure 6. Behavioral results from instar N1 to instar N5. Asterisks above the bars indicate significant differences from the expected results according to the analysis of standardized residuals: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. Effects of drone on H. halys behaviours from instar N1 to instar N5.

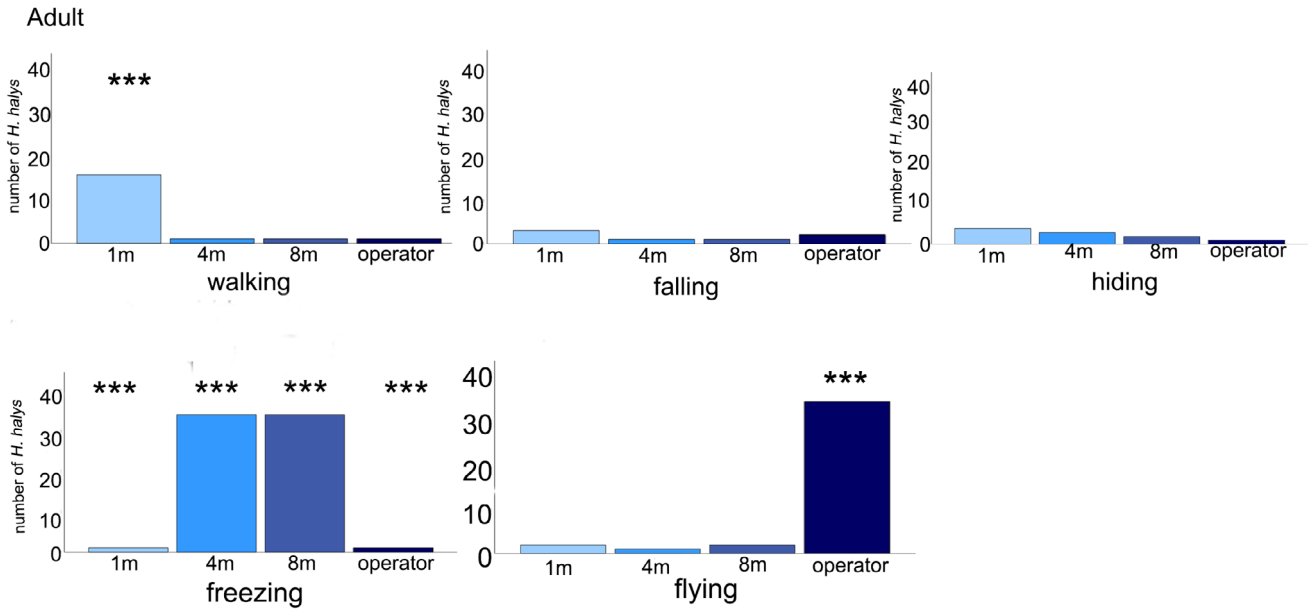

Figure 7. Effects of drone on adult behavior. Asterisks above the bars indicate significant differences from the expected according to the analysis of standardized residuals: *$P < 0.05$, ** $P < 0.01$, ***$P < 0.001$.

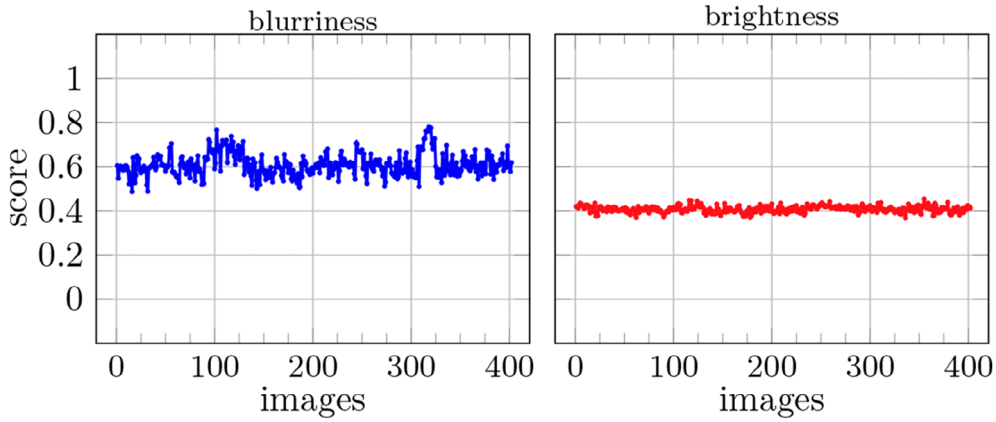

Figure 8. Evaluation of image blurriness (left) and brightness (right). The *x*-axis depicts the images and the *y*-axis their metrics score.

In detail, precision (P) describes the number of wrong predictions made, whereas recall (R) estimates the number of target objects missed during the detection. YOLO considers two parameters, IoU and confidence; namely, the percentage of overlap between a prediction and a target object, and the likelihood of a model in which the predicted object belongs to a certain category, respectively. To evaluate the confidence of the model's predictions, we rely on the $mAP_{0.5}$. For simplicity, the larger the $mAP_{0.5}$ value, the larger the confidence associated with the predictions. During the validation procedure, TL models perform better than those trained from scratch because of the learning capacity that stems from TL. Indeed, the models start from a certain recognition capacity, even though developed for a different task, and therefore receive a recognition boost. We then observe that the X model always outperforms the M model because, as the largest model available, the X model can handle the most complex functions of the YOLO family. However, although each curve stands out with its characteristics, all perform with a stable ascending trend reaching a plateau before the last epoch. This confirms that the parameters set for the training procedure are appropriate. Relying on the remaining 10% of unseen samples, we then assessed the models' detection abilities. For this purpose we consider, in addition to the previous metrics, the $mAP_{0.95}$, which reflects the impact of different IoUs upon the precision and recall curve.

In principle, looking at Table 2, the evaluation metrics indicate satisfactory performance, because each model demonstrates commendable precision (P) and recall (R). For instance, the X-scratch model reaches the highest precision score (100%), although registering the lowest detection rate (62%). Similarly, M-scratch exhibits notable precision (97%) but is again accompanied by an unbalanced recall score (67%). This imbalance between precision and recall represents a common behavior; in fact, models trained from scratch tend to be more conservative, with higher precision but lower recall. This effect stems from the fact that models trained from scratch, namely those with their initial weights set randomly, are susceptible to being trapped in local minima during the

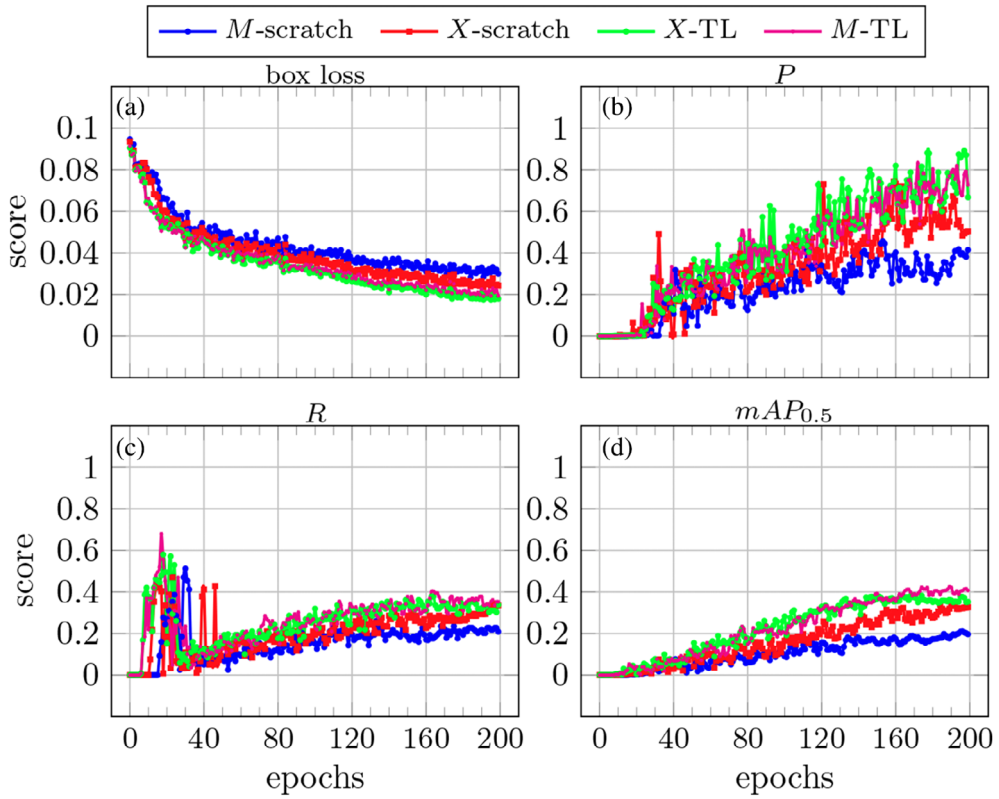

Figure 9. Training and validation performance. (a) Box loss depicts the ability to draw correctly bounding boxes. (b) Precision P captures the recognition ability. (c) Recall R measures how many objects were detected. (d) Mean average precision (mAP) harmonizes trends, providing a comprehensive performance snapshot for a detection task.

learning process. Eventually, a local minimum is reached, and for each further iteration the models then reinforce the previous knowledge, neglecting the possibility of generalizing their detection abilities. As a result, only a subset of metrics is optimized correctly. Models trained using TL consistently deliver the best performance across all metrics. In particular, they consistently outperform their counterparts trained from scratch across the metrics. Notably, X-TL achieves the highest overall metrics with a precision of 97% and a recall of 87% (Table 2). Once again, despite M-TL obtaining noteworthy performance, its X counterpart exhibits superior metrics (P = 94%, R = 83%). The reason for the outstanding performance is the introduction of TL during the learning process. TL was crucial to support generalization of the models' knowledge. Indeed, it achieves almost the same results as scratch models with respect to the precision score, improving the ability to find *H. halys*, namely the recall score. In addition, examining the mean average precision at IoU = 0.5 ($mAP_{0.5}$) reveals a limited gap with precision, suggesting a high level of confidence for every prediction. Conversely, examination of $mAP_{0.95}$ underscores the importance of the IoU threshold for predictions. The smaller value of $mAP_{0.95}$ with respect to $mAP_{0.5}$ reveals the confidence when the IoU threshold varies. Hence, we can state that the models predict with high confidence, although the intersection between predictions and the target objects sometimes appears limited. It should be noted that the result is more than acceptable in view of the limited size of *H. halys* in the average case. Although contouring tightly an *H. halys* instance is desirable, recognizing at least 50% of *H. halys* represents an outstanding milestone for an autonomous pest management system which relies on aerial images.

4 DISCUSSION

The results of the behavioral tests on the different *H. halys* instars showed that the hovering drone resulted in a reduction in disturbance compared with the presence of an operator. Specifically, the operator's presence triggered an increase in falling from leaves among all immature instars and an increase in flying among adults. These behaviors appear to be escape responses that are put in place to avoid a predator, suggesting that the operator may be perceived as a potential threat. When exposed to the hovering drone some of the N5 and most of the adults showed freezing, a behavior observed for the first time in *H. halys*. The response to the stimuli for adults only is likely due to the cryptic behavior of other stages. The thinner cuticle of young individuals compared with older immature stages and adults affects their protection against desiccation, resulting in the greater presence and activity of the younger stages in more sheltered portions of the plants, such as among leaves or below leaves and fruits. Freezing was only observed in the presence of wind at a speed ≥1 m/s, such as that produced by computer cooling fans or by the hovering drone (4 or 2 m/s at 4 or 8 m height, respectively). Freezing behavior has been described as a response exhibited by *Plautia stali* (Hemiptera, Pentatomidae) when exposed to artificially induced substrate vibrations at 150 and 500 Hz.[42] This behavior involves the insect stopping during antennation and grooming and remaining motionless. This freezing behavior was one of

Table 2. Model testing results

Model*	*H. halys* (HH) class			
	P†	R‡	$mAP_{0.5}$§	$mAP_{0.95}$§
X-scratch	1	0.62	0.85	0.43
M-scratch	0.96	0.67	0.84	0.33
X-TL	0.97	0.87	0.98	0.51
M-TL	0.94	0.83	0.92	0.42

*Models under consideration.
† Precision score, the model the recognition ability.
‡ Recall score, namely how many objects were detected.
§ Intuition of the predictions' quality with respect to confidence ($mAP_{0.5}$) and intersection with target objects ($mAP_{0.95}$).

several responses observed, along with squatting with bending all legs and walking, in reaction to the vibrations. The authors suggest that freezing behavior is a common response to artificial vibrations and may be a shared behavior among different insect species. According to Lee and Leskey,[43] wind speeds >0.75 m/s have an inhibitory effect compared with still air, drastically reducing the proportion of *H. halys* that take off. Therefore, it seems that freezing behavior is exhibited to counteract a wind that is too strong. Substrate-born vibrations are typically used by *H. halys* during intraspecific communication, especially courtship and mating[44]; however, freezing behavior has never been observed when bugs are exposed to substrate vibrations produced by electromagnetic shakers ranging at 72 to 152 Hz.[45] In any case, according to the behavioral experiments, the hovering drone is much less disruptive to the bugs than the presence of an operator. The drone-induced freezing behavior allows for improved detection through drone monitoring and is therefore functional for the project. In fact, the bugs on the plants do not fly away, which makes estimation of the insects that are actually present more reliable than with manual monitoring, which makes them flee. Furthermore, the fact that the bugs remain motionless makes it possible to capture high-resolution images more easily.

Image analysis demonstrated significant results supporting the use of cameras in the field for monitoring in different environmental situations and significant results for the YOLO-trained models. Each model showed very strong performance with precision up to 100% and recall up to 87%. The X-TL model showed the best performance, identifying the greatest number of bugs with the fewest errors. This difference compared with other models could be attributed to the model's size (the number of layers and nodes in the network, which characterize better learning functions) and the use of TL, which maybe accelerates the convergence of metrics. These results establish a drone imaging mechanism as a state-of-art method of *H. halys* detection.

5 CONCLUSIONS

This study demonstrates the potential use of UAVs for automated monitoring, particularly detecting the mobile stages of invasive species like *H. halys*. Implementation of UAVs and AI could optimize insect monitoring times by reducing travel time and providing a ready-to-use solution to assess the infestation rate of monitored areas in real time; in contrast to traditional time-consuming monitoring methods. These techniques provide decision support for farmers and researchers in agroecosystem management. The ability to capture high-altitude images with good resolution makes this system suitable in different agricultural or forestry contexts. Furthermore, the impressive results obtained using our AI system demonstrate the potential for autonomous identification and quantification of additional target species and for the assessment of indirect crop damage caused by other pest species, including leaf and/or fruit deformation. Further research could extend the use of AI models trained to other pests and plant pathogens. In addition, the designed protocol, which involved the use of drones, could be further optimized to allow for monitoring with a greater frequency and coverage. Finally, autonomous detection data could be integrated with environmental data to develop an abundance and distribution model of the pest.

ACKNOWLEDGEMENTS

HALY.ID is part of ERA-NET Cofund ICT-AGRI-FOOD, with funding provided by MIPAAF and co-funding by the European Union's Horizon 2020 research and innovation program, Grant Agreement number 862665ERA-NET ICT-AGRI-FOOD (HALY-ID 862671).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Hassanalian M and Abdelkefi A, Classifications, applications, and design challenges of drones: a review. *Prog Aerosp Sci* **91**:99–131 (2017).
2 Tiberiu PB, Gheorghe FB and Constantin B, The use of drones in forestry. *J Environ Eng Sci* **5**:11 (2016).
3 Donnarumma L, D'Argenio A, Sandulli R, Russo GF and Chemello R, Unmanned aerial vehicle technology to assess the state of threatened biogenic formations: the vermetid reefs of mediterranean intertidal rocky coasts. *Estuar Coast Shelf Sci* **251**:107228 (2021).
4 Vermeulen C, Lejeune P, Lisein J, Sawadogo P and Bouché P, Unmanned aerial survey of elephans. *PloS One* **8**:e54700 (2013).
5 Michez A, Broset S and Lejeune P, Ears in the sky: potential of drones for the bioacoustic monitoring of birds and bats. *Drones* **5**:9 (2021).
6 Koski WR, Allen T, Ireland D, Buck G, Smith PR, Macrander AM *et al.*, Evaluation of an unmanned airborne system for monitoring marine mammals. *Aquat Mamm* **35**:347–357 (2009).

7 Lake EC, David AS, Spencer TM, Wilhelm VL, Barnett TW, Abdel-Kader AA *et al*., First drone releases of the biological control agent *Neomusotima conspurcatalis* on old World climbing fern. *Biocontrol Sci and Technol* **31**:97–106 (2021).
8 Zhang H, Wang L, Tian T and Yin J, A review of unmanned aerial vehicle low altitude remote sensing (UAV-LARS) use in agricultural monitoring in China. *Remote Sens (Basel)* **13**:1221 (2021).
9 Popescu D, Ichim L and Stoican F, Orchard monitoring based on unmanned aerial vehicles and image processing by artificial neural networks: a systematic review. *Front Plant Sci* **14** (2023). https://doi.org/10.3389/fpls.2023.1237695.
10 Park YL, Cho JR, Lee GS and Seo BY, Detection of *Monema flavescens* (lepidoptera: Limacodidae) cocoons using small unmanned aircraft system. *J Econ Entomol* **114**:1927–1933 (2021).
11 Maistrello L, Dioli P, Dutto M, Volani S, Pasquali S and Gilioli G, Tracking the spread of sneaking aliens by integrating crowdsourcing and spatial modeling: the Italian invasion of *Halyomorpha halys*. *Bioscience* **68**:979–989 (2018).
12 Maistrello L, Vaccari G, Caruso S, Costi E, Bortolini S, Macavei L *et al*., Monitoring of the invasive *Halyomorpha halys*, a new key pest of fruit orchards in northern Italy. *J Pest Sci* **90**:1231–1244 (2017).
13 CSO Italy, Estimation of damage from brown marmorated stink bug and plant pathologies related to climate change. (2020) Available: http://www.csoservizi.com [20 December 2023].
14 Kirkpatrick DM, Acebes-Doria AL, Rice KB, Short BD, Adams CG, Gut LJ *et al*., Estimating monitoring trap plume reach and trapping area for nymphal and adult *Halyomorpha halys* (Hemiptera: Pentatomidae) in crop and non-crop habitats. *EnvironEntomol* **48**:1104–1112 (2019).
15 Morrison WR, Lee DH, Short BD, Khrimian A and Leskey TC, Establishing the behavioral basis for an attract-and-kill strategy to manage the invasive *Halyomorpha halys* in apple orchards. *Journal of Pest Sci* **89**:81–96 (2016).
16 Sargent C, Martinson HM and Raupp MJ, Trap placement may affect location of Brown marmorated stink bug (Hemiptera: Pentatomidae) and increase injury to tomato fruits in home gardens. *Environ Entomol* **43**:432–438 (2014).
17 Betti Sorbelli F, Corò F, Das SK, Palazzetti L and Pinotti MC, Drone-based optimal and heuristic orienteering algorithms towards bug detection in orchards, in *18th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. Marina del Rey, Los Angeles, CA, USA, pp. 117–121 (2022).
18 De Oca AM and Flores G, The agriq: a low-cost unmanned aerialsystem for precision agriculture. *Expert Systems* **18**:115163 (2021).
19 Xie C, Zhang J, Li R, Li J, Hong P, Xia J *et al*., Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning. *Comput. Electron. Agric* **119**:123–132 (2015).
20 Wen C and Guyer D, Image-based orchard insect automated identification and classification method. *Comput Electron Agric* **89**:110–115 (2012).
21 Li W, Wang D, Li M, Gao Y, Wu J and Yang X, Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse. *Comput. Electron. Agric* **183**:106048 (2021).
22 Valan M, Makonyi K, Maki A, Vondracek D and Ronquist F, Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. *Syst Biol* **68**: 876–889 (2019).
23 Espinoza K, Valera DL, Torres JA, Lopez A and Molina Aiz FD, Combination of image processing and artificial neural networks as a novel approach for the identification of Bemisia tabaci and Frankliniella occidentalis on sticky traps in greenhouse agriculture. *Comput. Electron. Agric* **127**:495–505 (2016).
24 Betti Sorbelli F, Palazzetti L and Pinotti MC, Preliminary results for *Halyomorpha halys* monitoring relying on a custom dataset, in *2023 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)*. IEEE, Pisa, Italy, pp. 363–436 (2023).
25 Dinca A, Popescu D, Pinotti CM, Ichim L, Palazzetti L and Angelescu N, Halyo*morpha halys* detection in orchard from UAV images using convolutional neural networks, in *Advances in computational intelligence. IWANN 2023. lecture notes in computer science*, Vol. **14135**, ed. by Rojas I, Joya G and Catala A. Springer, Cham (2023).
26 Khrimian A, Zhang A, Weber DC, Ho HY, Aldrich JR, Vermillion KE *et al*., Discovery of the aggregation pheromone of the brown marmorated stink bug (*Halyomorpha halys*) through the creation of stereoisomeric libraries of 1-bisabolen-3-ols. *J Nat Prod* **77**:1708–1717 (2014).
27 Weber DC, Leskey TC, Walsh GC and Khrimian A, Synergy of aggregation pheromone with methyl (E, E, Z)-2, 4, 6-decatrienoate in attraction of *Halyomorpha halys* (Hemiptera: Pentatomidae). *J Econ Entomol* **107**:1061–1068 (2014).
28 HALY.ID, *Halyomorpha halys* Identification: Innovative ICT tools for targeted monitoring and sustainable management of the brown marmorated stink bug and other pests (2023). https://www.haly-id.eu/about/. Accessed 13 Nov 2023.
29 Ekaso D, Nex F and Kerle N, Accuracy assessment of real-time kinematics (RTK) measurements on unmanned aerial vehicles (UAV) for direct geo-referencing. *Geo-spatial information science* **23**:165–181 (2020).
30 Costi E, Haye T and Maistrello L, Surveying native egg parasitoids and predators of the invasive *Halyomorpha halys* in northern Italy. *J. Appl. Entomol.* **143**:299–307 (2019).
31 Schifani E, Giannetti D, Costi E, Franconi G, Campostrini A, Maistrello L *et al*., Interactions between egg parasitoids and predatory ants for the biocontrol of the invasive brown marmorated stink bug *Halyomorpha halys*. *J Appl Entomol* **147**:868–874 (2023). https://doi.org/10.1111/jen.13179.
32 Humphreys RK and Ruxton GD, A review of thanatosis (death feigning) as an anti-predator behaviour. *Behav Ecol Sociobiol* **72**:1–16 (2018).
33 Betti Sorbelli F, Corò F, Das SK, Di Bella E, Maistrello L, Palazzetti L *et al*., A drone-based application for scouting *Halyomorpha halys* bugs in orchards with multifunctional nets, in *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, Pisa, Italy (2022).
34 Betti Sorbelli F, Palazzetti L and Pinotti CM, YOLO-based detection of *Halyomorpha halys* in orchards using RGB cameras and drones. *Comput Electron Agri* **213**:108228 (2023).
35 Narvekar ND and Karam LJ, A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *IEEE Trans Image Process* **20**:2678–2683 (2011).
36 Ware C, Chapter three—lightness, brightness, contrast, and constancy, in *Information Visualization (Fourth Edition)*, Fourth edn, ed. by Ware C. Interactive Technologies, Morgan Kaufmann, Netherlands, pp. 69–94 (2021).
37 Redmon J, Divvala S, Girshick R and Farhadi A, You only look once: unified, real-time object detection, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. **2016**. Las Vegas, NV, USA, pp. 779–788 (2016).
38 Jocher G, Chaurasia A, Stoken A, Borovec J, Kwon Y, Fang J *et al*., ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpuand openvino export and inference Zenodo, (2022).
39 Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H *et al*., A comprehensive survey on transfer learning. *Proc IEEE* **109**:43–76 (2020).
40 Almstedt L, Baltieri D, Betti Sorbelli F, Cattozzi D, Giannetti D, Kargar A *et al*., Technological innovations in agriculture for scouting *Halyomorpha halys* in orchards, in *19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*. IEEE, Paphos, Cyprus, pp. 702–709 (2023).
41 Uechi N and Takanashi T, Behavioral control of *Plautia stali* (Hemiptera: Pentatomidae) using vibrations, and its application in pest management. *Jpn J Appl Entomol* **65**:13–20 (2021).
42 Lee DH and Leskey TC, Flight behavior of foraging and overwintering brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Pentatomidae). *Bull Entomol Res* **105**:566–573 (2015).
43 Polajnar J, Maistrello L, Bertarella A and Mazzoni V, Vibrational communication of the brown marmorated stink bug (*Halyomorpha halys*). *Physiological Entomology* **41**:249–259 (2016).
44 Mazzoni V, Polajnar J, Baldini M, Rossi Stacconi MV, Anfora G, Guidetti R *et al*., Use of substrate-borne vibrational signals to attract the brown marmorated stink bug. *Halyomorpha halys J Pest Sci* **90**:1219–1229 (2017).
45 Caorsi V, Cornara D, Wells KE, Moser D, Berardo A, Miselli R *et al*., Design of ideal vibrational signals for stinkbug male attraction through vibrotaxis experiments. *Pest Manag Sci* **77**:5498–5508 (2021).